4-2019

Exploring and visualizing household electricity consumption patterns in Singapore: A geospatial analytics approach

Yong Ying TAN
Singapore Management University, yy.tan.2017@mitb.smu.edu.sg

Tin Seong KAM
Singapore Management University, tskam@smu.edu.sg

DOI: https://doi.org/10.1007/978-3-030-15742-5_74

Exploring and Visualizing Household Electricity Consumption Patterns in Singapore: A Geospatial Analytics Approach

[1]Joanne Tan Yong Ying and [2]Kam Tin Seong

Singapore Management University, 80 Stamford Road Singapore 178902, Singapore
[1]yy.tan.2017@mitb.smu.edu.sg
[2]tskam@smu.edu.sg

Abstract. Despite being a small country-state, electricity consumption in Singapore is said to be non-homogeneous, as exploratory data analysis showed that the distributions of electricity consumption differ across and within administrative boundaries and dwelling types. Local indicators of spatial association (LISA) were calculated for public housing postal codes using June 2016 data to discover local clusters of households based on electricity consumption patterns. A detailed walkthrough of the analytical process is outlined to describe the R packages and framework used in the R environment. The LISA results are visualized on three levels: country level, regional level and planning subzone level. At all levels we observe that households do cluster together based on their electricity consumption. By faceting the visualizations by dwelling type, electricity consumption of planning subzones can be said to fall under one of these three profiles: low-consumption subzone, high-consumption subzone and mixed-consumption subzone. These categories describe how consumption differs across different dwelling types in the same postal code (HDB block). LISA visualizations can guide electricity retailers to make informed business decisions, such as the geographical zones to enter, and the variety and pricing of plans to offer to consumers.

Keywords: Electricity Consumption, Exploratory Spatial Data Analysis, Spatial Autocorrelation.

1 Introduction

Currently, electricity consumption in Singapore is closely monitored by the Energy Market Authority (EMA) and reported on annually through the Singapore Energy Statistics publication [1]. Despite being a small country-state, there are observed differences in electricity consumption across geographical areas and dwelling types (1-room and 2-room flats, 3-room flats, 4-room flats, 5-room and executive flats) which could be attributed to demographic factors such as household size, household income or proportion of economically active population in the region.

Looking at the average monthly household electricity consumption by dwelling type in Table 3.4 of Singapore Energy Statistics report, it is obvious that households in bigger dwelling types have higher consumption [1]. However, it becomes difficult to compare across planning zones in a tabular format in Table 3.5 [1]. Thus, the author downloaded the data to perform an exploratory data analysis to discover if electricity usage is indeed heterogeneous across dwelling types or geographical areas.

Figure 1 shows the distribution of consumption by postal codes in each planning area, faceted by dwelling type and sorted by median consumption. The results are consistent with those reported by EMA in that bigger dwelling types correspond to higher electricity usage. We also see that consumptions are different across planning areas of each dwelling type, proving that electricity usage is heterogeneous across both dwelling types and geographical areas. On top of that, we see that within each planning area, distributions are wide and most of them have outliers, which debunks generalizations that electricity usage is homogeneous within an administrative area.

Motivated by the differences observed at multiple facets: across dwelling types, across planning areas and within planning areas, this study will explore the application of local indicators of spatial association (LISA) on electricity consumption data in Singapore to answer two questions: where are the outliers located at? And do households of high electricity usage cluster together spatially, or are they randomly distributed throughout Singapore? As we have the advantage of possessing highly disaggregated data at the postal code level, we will explore spatial autocorrelation on point features, which is different from most LISA applications that explore spatial autocorrelation on polygon features.

In this paper, we will use data provided by EMA [2] in the month of June 2016 to analyze spatial autocorrelation of electricity consumption among Singapore's public housing postal codes using the local Moran's I statistic. Then, we will interpret the statistic into four quadrants and map the results on three levels. The maps are faceted by dwelling type, so that we can identify local clusters of electricity consumption patterns and observe how the clusters are similar or dissimilar across dwelling types in the same administrative area. In the end, we come up with three hypotheses for future work.

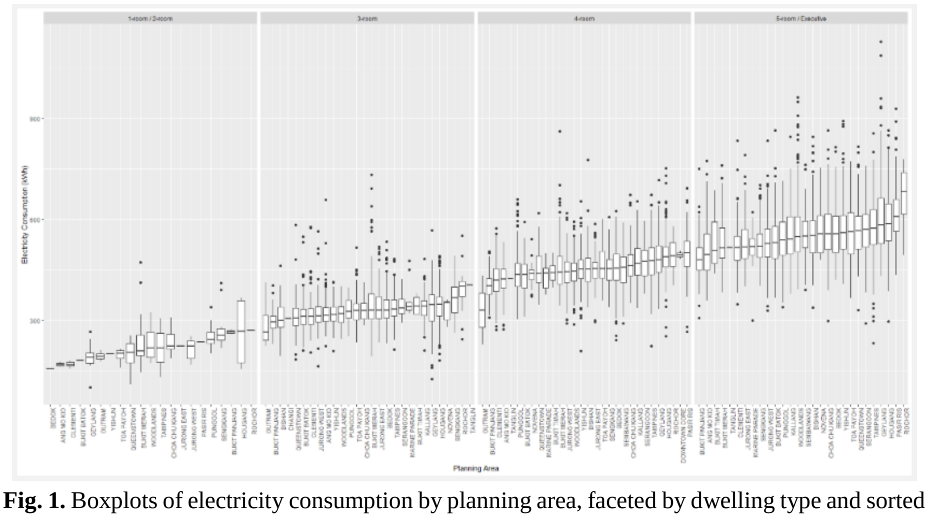

Fig. 1. Boxplots of electricity consumption by planning area, faceted by dwelling type and sorted by median electricity consumption.

2 Literature Review

2.1 Analysis and Modeling of Energy Consumption outside Singapore

Urquizo, Calderon and James analyzed spatial patterns of Annual Energy Consumption Intensity (AECI) in South Heaton, Westgate and Castle districts using hotspot maps [3]. They investigated Houses in Multiple Occupation (HMOs) which serve low income people, students and migrants who do not have much money to spend on accommodation. As HMOs are rented on a room-by-room basis, typical measures of density such as 'population density' or 'dwelling density' may not be suitable for modeling energy consumption in these dwellings. Thus, 'space per person' is proposed as a new measure, which is defined as "the dwelling size of a habitable unit divided by household size".

Urquizo also developed a framework to estimate energy consumption in sub-city areas on three levels: district level, neighborhood level and a retrofitted model from single dwellings to community level [4]. The study found that national or regional indices often do not capture important factors in modeling energy consumption that exist at the local area level. For example, the neighborhood model had over- or underestimated energy consumption in all the three districts studied, because the assumptions made for each district/neighborhood were unable to account for its heterogeneity in building types, resident background and energy systems. Therefore, the paper concluded that local area characteristics should be included in energy modeling to produce more robust estimates that can influence policy decisions more effectively.

Yang, S. et. al. used a city-wide household survey to collect and map data on carbon emissions in Shanghai [5]. It also found that carbon emissions are not homogeneous throughout Shanghai, as the spatial pattern of carbon emissions by households can be described as donut-shaped: low in the urban center, high in suburban areas, and low again in rural areas. Factors that were found to influence household emissions include car ownership, type/size/age of dwelling and household income, all of which cannot be easily described or detected at the highly aggregated level.

2.2 Analysis and Modeling of Energy Consumption in Singapore

There are also some studies done on electricity consumption in Singapore. Agarwal, Satyanarain, Sing, and Vollmer [6] investigated how construction activities led to an increase in electricity consumption as residents sought to reduce the impact of noise pollution by closing their windows and using air conditioners. They found that electricity consumption does not rebound back to pre-construction levels after the construction works were completed, which could be a significant finding for stakeholders interested in reducing energy consumption in Singapore.

Loi and Ng [7] found that residential users in Singapore are more sensitive to price decreases than price increases of electricity, which implies it may be difficult to induce energy conservation via increasing electricity prices. Other motivators, such as education or the creation of green champions in communities, should be used to promote energy conservation instead.

Luo and Ukil simulated electrical load profiles of dwelling types in Singapore based on a probabilistic model of when an electrical appliance will be activated and the average time of appliance usage daily, before the load profiles were successfully verified against measurements of actual electricity consumption at campus housings in National Technological University [8]. Their work successfully established a bottom-up approach of simulating households' electricity consumption based on one household or one electric appliance, which could be helpful for future studies on the smart grid.

3 Methodology

3.1 Local Indicators of Spatial Association (LISA)

Local indicators of spatial association were introduced by Anselin in 1995 [9] to identify local spatial clusters which may not be picked up by global indicators or uncover local spatial trends that are opposite of global spatial patterns. The null hypothesis assumes that there is no autocorrelation between attribute values in neighboring features, while the alternative hypothesis states that the neighboring features are spatially autocorrelated and therefore said to be spatially clustered. A positive statistic value denotes that the neighboring features are similar, while a negative statistic denotes dissimilarity.

Taking the 2016 data on electricity consumption from Energy Market Authority, a Local Moran's I statistic is calculated for each postal code each month using equation (1). x_i refers to the electricity consumption of postal code i expressed in kilowatt hour (kWh), $w_{i,j}$ refers to the spatial weight between postal codes i and j, $\bar{X}$ refers to the national average electricity consumption and n refers to the total number of target postal codes included for comparison with the origin postal code. Since a rule of thumb is to have at least 30 features in each calculation to achieve reliable results [10], $n = 30$ in this research.

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j} (x_j - \bar{X}) \tag{1}$$

Based on the point pattern distribution of the raw dataset, we ran 1000 Monte Carlo simulations under the assumption that there is no spatial pattern in electricity consumption of households. After getting a Local Moran's I statistic for each postal code, its statistic is interpreted by calculating which of the five categories it falls into based on their electricity consumption figures relative to Singapore's average value, statistic value and P-value:

1. Insignificant: Even though I have a local Moran's I statistic, it is not statistically significant at the p-value of 0.1.
2. Low-low: If my electricity consumption is lower than the national average and my neighbors' consumptions are as low or lower, I am a low-low point.
3. Low-high: If my electricity consumption is lower than the national average but my neighbors' consumptions are relatively higher than me, I am a low outlier among high points.

4. High-low: If my electricity consumption is higher than the national average but my neighbors' consumptions are not as high as me, I am a high outlier among low points.
5. High-high: If my electricity consumption is higher than national average and my neighbor's consumptions are as high/higher than me, I am a high-high point.

3.2 Implementing LISA Analysis in R Environment

R-Markdown. The R Markdown framework was chosen to document the analytical process and methods. It is a tool for good coding practices as the R-Markdown document is easily distributed, reproducible and comes with a well-rounded syntax for documentation [11]. It also enables users to write code in chunks, which is excellent for debugging and ensuring readability. The R Markdown document makes it easy for the author to edit only the relevant parameters to generate new output on an ad-hob basis without having to redo the entire code chunk.

LISA Methods in R. The LISA analysis is powered by the 'spdep' package which provides methods for analyzing spatial dependence [12]. Firstly, for each postal code in the dataset, a list of 30 nearest neighbors is generated using the '*knearneigh'* and '*knn2b*' functions. Then, distances were calculated between the origin point and the 30 neighbors using '*nbdists*', before inverse distance weights are supplied to the neighbors using '*nb2listw*'. Finally, the local Moran's I statistic for each postal code is calculated using the '*localmoran*' function, before its quadrant is calculated based on its local Moran's I statistic and electricity consumption figures relative to its neighbors and to Singapore.

LISA Visualizations. Finally, the interpreted LISA results are plotted on faceted visualizations of maps using the 'tmap' package [13]. Methods for generating visualizations in 'tmap' follow the same grammar as 'ggplot2', meaning a base 'tmap' object is first defined using '*tm_shape(<spatial object>)'* before multiple methods are "added" on top of the base object to fine-tune the aesthetics such as color, transparency and shapes. This feature makes it much more desirable to plot maps using 'tmap' rather than base R functions which produce plots that look less refined and are difficult for the researcher to modify when one wants to try adding or removing elements from the visualization.

4 Findings and Discussion

4.1 Overview of Clusters in Singapore

Figure 2 shows a plot of postal codes color-coded by their respective quadrants in June 2016. The plot is faceted by Dwelling Type to show differences in distributions across different types of housing. 1-room/2-room flats do not show obvious clusters in the plot due to the low number of postal codes (Table 1). Across all dwelling types, clear hot clusters and cold clusters can be observed in different parts of Singapore, which implies that households of similar consumption patterns indeed cluster together geographically. The East region has many clusters of households with high electricity usage, while the West region has many clusters of households with low electricity usage.

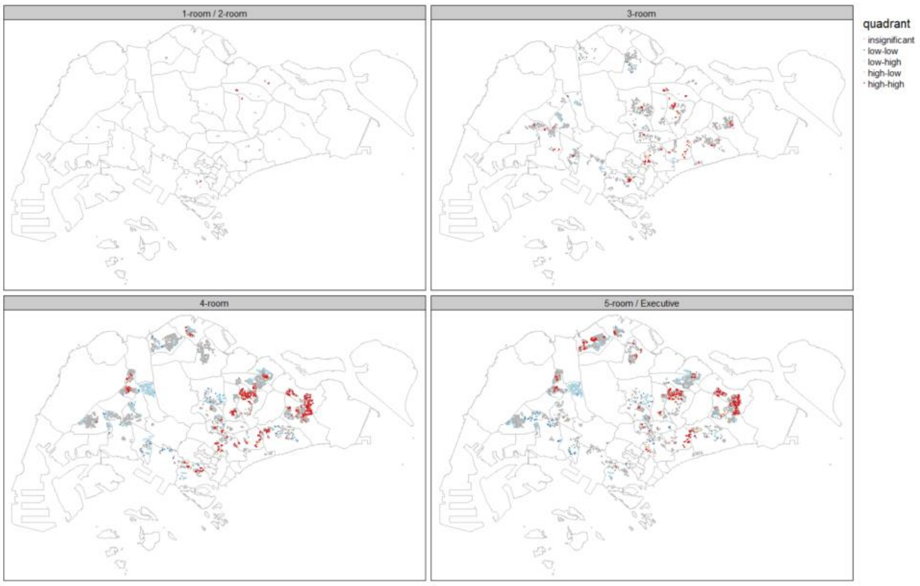

Fig. 2. Postal codes in Singapore, color-coded by their respective quadrants in June 2016. Plot is faceted by dwelling type.

Table 1. Number and proportion of units for each dwelling type

	1-room/2-room	3-room	4-room	5-room/Executive Flats
Insignificant	97 (0.78%)	1646 (13.30%)	3603 (29.12%)	3395 (27.44%)
Low-low	0 (0%)	14 (0.11%)	75 (0.61%)	97 (0.78%)
Low-high	11 (0.09%)	168 (1.36%)	789 (6.38%)	675 (5.46%)
High-low	2 (0.02%)	14 (0.11%)	45 (0.36%)	49 (0.40%)
High-high	10 (0.08%)	173 (1.40%)	801 (6.47%)	708 (5.72%)
Total	**120 (0.97%)**	**2015 (16.29%)**	**5313 (42.94%)**	**4924 (39.80%)**

4.2 Overview of Clusters at Planning Area Level

When we zoom down into the planning area level, we see two scenarios. The first scenario is when the HDB blocks exhibit similar consumption patterns across dwelling types, such as that observed in the North-East region (Figure 3). There is a cluster of households (denoted with a solid outline circle) that exhibit high consumption patterns throughout, while there is another cluster of households (denoted with a dashed outline circle) that exhibit low consumption patterns throughout.

The second scenario is when the HDB blocks exhibit different consumption patterns across dwelling types, such as that observed in the North region (Figure 4). In many subzones in this region, there are clusters of HDB blocks that exhibit low consumption among 4-room households but high consumption among 5-room households. We drilled down further into several planning subzones within these regions to further investigate differences across dwelling types.

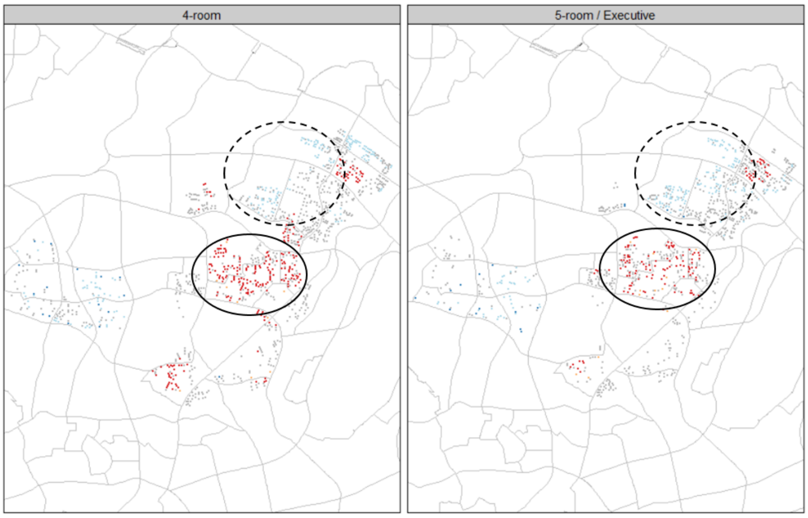

Fig. 3. Postal codes in North-East region. Plot is faceted by dwelling type and cropped to show differences between 4-room and 5-room households.

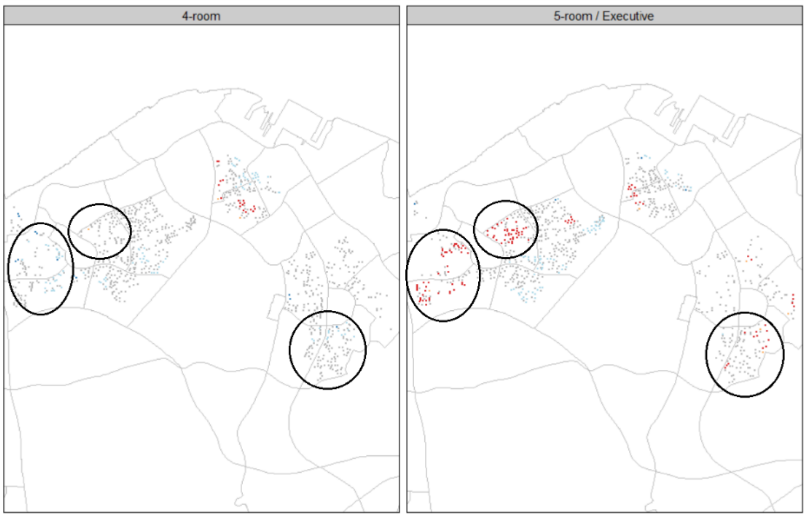

Fig. 4. Postal codes in North region. Plot is faceted by dwelling type and cropped to show differences between 4-room and 5-room households.

4.3 Overview of Clusters at Planning Subzone Level

Planning subzones are the lowest level of administrative boundaries in Singapore. Generally at the subzone level, three types of phenomenon can be observed:

1. Low consumption subzone: Households across all dwelling types in this subzone are low-low points.
2. High consumption subzone: Households across all dwelling types in this subzone are high-high points.
3. Mixed-consumption subzone: Households in this subzone exhibit different consumption habits at different dwelling types (for example 3-room households may be low-consumption households, while 4-room households may be high-consumption households in the same subzone).

Subzones in the North Region. In Figure 4 we saw there were distinct clusters of HDB blocks that had high-consumption 5-room households but low-consumption 4-room households. Examples of such mixed-consumption subzones include Midview (Figure 5), Woodgrove (Figure 6) and Yishun South. Out of three dwelling types in Midview, only the 5-room/executive flat households have high electricity usage, but even within this dwelling type, we see there are two distinct clusters: one hot cluster at the northern part and one cold cluster at the southern part. We also identified one block which showed different consumption characteristics across dwelling types: it is a high consumption outlier among 4-room flats but is a member of a high consumption cluster among 5-room flats.

In Woodgrove (Figure 6), differences across dwelling types are more pronounced: generally, 4-room households are of low consumption while 5-room households are of high consumption. Even among 4-room households, consumption patterns are not homogeneous: there is a cluster of low outliers among high consumption households, while there is another cluster of low consumption household.

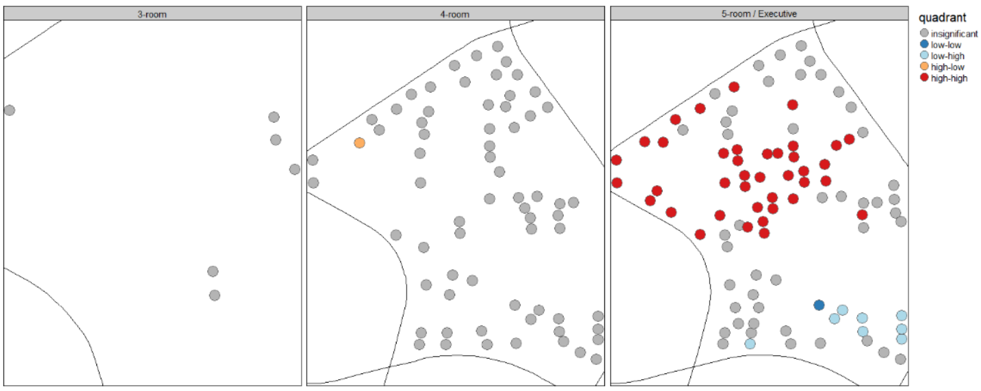

Fig. 5. Midview Planning Subzone located in the North region.

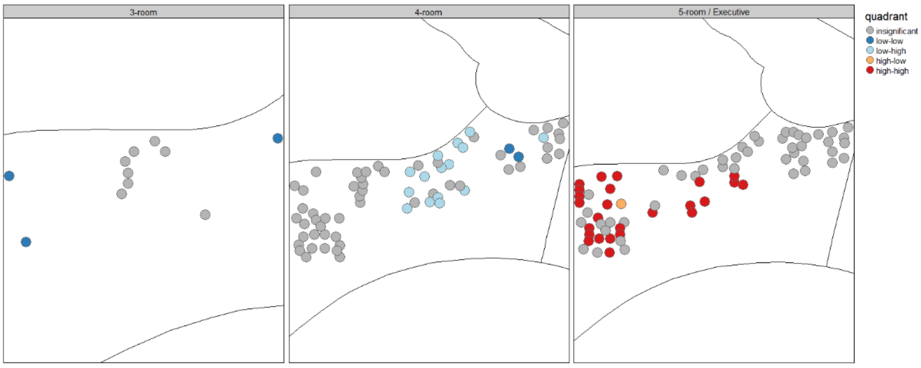

Fig. 6. Woodgrove Planning Subzone located in the North region.

Subzones in the North-East Region. In the North-East region, there are few mixed-consumption subzones, and most subzones are either purely high-consumption or low-consumption. Examples of high-consumption subzones are Hougang West (Figure 7), Hougang East and Trafalgar, while examples of low-consumption subzones are Anchorvale (Figure 8) and Matilda.

Fig. 7. Hougang West Planning Subzone located in the North-East region.

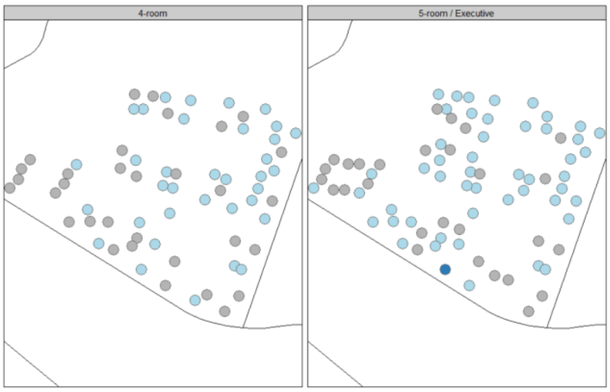

Fig. 8. Anchorvale Planning Subzone located in the North-East region.

4.4 Overall Discussion of Results

These results highlight the value of analyzing spatial data at lower levels of visualization, and more importantly the value of increasingly disaggregated data. In Section 4.1, the country overview only enables us to make general conclusions on the regional level. As we go deeper, we start to reveal differences within one area that could not be observed in the macro level: electricity consumption is not necessarily homogeneous within each dwelling type or administrative area. In fact, differences have been identified at every level of visualization, which prompts us to reflect that any generalizations, forecasting models or time-series predictions that do not account for spatial information will probably not perform well, and we should rethink what other disaggregated factors could influence electricity consumption habits.

4.5 Relevance to Singapore Electricity Market

Our work will be increasingly relevant to electricity retailers in Singapore as the country is moving towards the concept of an "Open Electricity Market" (OEM). Previously, there were no "electricity retailers" in Singapore's electricity industry structure, and non-contestable consumers (households and small businesses) could only purchase electricity from SP Group at a regulated tariff. Starting 1 April 2018, consumers in selected zones of Singapore could choose to purchase electricity from one of the 13 retailers with a price plan [14]. Since the OEM is launched by geographical zones progressively, retailers can choose the zones they want to enter and the plans to offer in each zone. However, only three types of plans exist as of December 2018: i) a fixed rate per kWh; ii) a fixed percentage discount off the SP Group tariff; and iii) a higher fixed rate during peak hours and a lower fixed rate during off-peak hours to appeal to the "night owls" (refer to [15] for current list of retailers and plans). No retailers had offered plans to differentiate between high- or low-usage consumers. Here, our visualizations can help retailers to: i) decide on the zones to enter and ii) identify the different consumer groups by usage levels in each zone, thereby enable them to devise and offer a wider range of plans to cater to their different demands and improve market strategy.

5 Conclusion

First, the paper established that electricity usage patterns are not homogeneous throughout Singapore. Then, the LISA statistic is calculated for each postal code and interpreted as one of the four quadrants (low-low, low-high, high-low, high-high), before the postal codes are color-coded and plotted on faceted map visualizations to uncover similarities and differences in electricity usage among households across geographical boundaries and dwelling types. Visualizations reveal that consumption levels are not homogeneous both within and across geographical boundaries and dwelling types, thus we should consider testing the influence of more disaggregated factors on electricity consumption and see if they may explain the differences better than aggregated factors. Electricity retailers in Singapore can leverage on these visualizations to guide their business decisions on the zones to enter, and the variety and pricing of plans to offer.

References

1. Singapore Energy Statistics. https://www.ema.gov.sg/cmsmedia/Publications_and_Statistics/Publications/ses/2017/downloads/SES2017_Chapter_1_to_9.pdf
2. EMA : Statistics. https://www.ema.gov.sg/statistics.aspx
3. Urquizo J, Calderon C, James P (2016) A spatial perspective of the domestic energy consumption intensity patterns in sub-city areas. A case study from the United Kingdom. In: 2016 IEEE Ecuador Technical Chapters Meeting (ETCM). IEEE Press, New York, pp 1-7. doi: 10.1109/ETCM.2016.7750848
4. Urquizo J (2015) A spatial model for domestic end-use energy diagnostic and support of energy efficiency policy to reduce fuel poverty in UK (2015). http://proceedings.esri.com/library/userconf/proc15/papers/606_395.pdf
5. Yang S, Wang C, Lo K, Wang M, Liu L (2015) Quantifying and Mapping Spatial Variability of Shanghai Household Carbon Footprints. In: Frontiers in Energy, vol 9, no 1. Higher Education Press and Springer Verlag Berlin Heidelberg, pp 115–124. doi:10.1007/s11708-015-0348-8
6. Agarwal S, Satyanarim R, Sing T F, Vollmer D (2016) Effects of Construction Activities on Residential Electricity Consumption: Evidence from Singapore's Public Housing Estates. In: Energy Economics, vol 55. Elsevier, pp 101–111. doi:10.1016/j.eneco.2016.01.010
7. T S A Loi, J L Ng (2018) Analysing Households' Responsiveness towards Socio-Economic Determinants of Residential Electricity Consumption in Singapore. In: Energy Policy, vol 112. Elsevier, pp 415–426. doi:10.1016/j.enpol.2017.09.052
8. Luo C, Ukil A (2015) Modeling and Validation of Electrical Load Profiling in Residential Buildings in Singapore. In: IEEE Transactions on Power Systems, vol 30, no 5. IEEE Press, New York, pp 2800–2809. doi:10.1109/TPWRS.2014.2367509
9. Anselin L (1995) Local Indicators of Spatial Association—LISA. In: Geographical Analysis, vol. 27, no. 2. Wiley, pp 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x
10. How Cluster and Outlier Analysis (Anselin Local Moran's I) works. http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-cluster-and-outlier-analysis-anselin-local-m.htm
11. R Markdown Cheat Sheet. https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf
12. CRAN – Package spdep. https://cran.r-project.org/package=spdep
13. CRAN – Package tmap. https://cran.r-project.org/web/packages/tmap/index.html
14. EMA : Overview of Electricity Market. https://www.ema.gov.sg/electricity_market_overview.aspx
15. List of Retailers. https://www.openelectricitymarket.sg/residential/list-of-retailers